Deep learning based short-term total cloud cover forecasting.

BANDARA, I., ZHANG, L. and MISTRY, K.

2022

Deep Learning Based Short-Term Total Cloud Cover Forecasting

Ishara Bandara
School of Computing
Robert Gordon University
Aberdeen, United Kingdom
isharabnd@gmail.com

Li Zhang
Department of Computer Science
Royal Holloway, University of London
Egham, United Kingdom
li.zhang@rhul.ac.uk

Kamlesh Mistry
Computer and Information Sciences
Northumbria University
Newcastle upon Tyne, United Kingdom
k.mistry@northumbria.ac.uk

***Abstract*—In this research, we conduct deep learning based Total Cloud Cover (TCC) forecasting using satellite images. The proposed system employs the Otsu's method for cloud segmentation and Long Short-Term Memory (LSTM) variant models for TCC prediction. Specifically, a region-based Otsu's method is used to segment clouds from satellite images. A time-series dataset is generated using the TCC information extracted from each image in image sequences using a new feature extraction method. The generated time series data are subsequently used to train several LSTM variant models, i.e. LSTM, bi-directional LSTM and Convolutional Neural Network (CNN)-LSTM, for future TCC forecasting. Our approach achieves impressive average RMSE scores with multi-step forecasting, i.e. 0.0543 and 0.0823, with respect to both the first half of daytime and full daytime TCC forecasting on a given day, using the generated dataset.**

***Index Terms*—Long Short-Term Memory, Deep Learning, Total Cloud Cover, Time-series Forecasting, and Satellite Imaging**

I. Introduction

As the world focuses more on minimizing the use of fossil fuels, many have moved into renewable energy sources like photovoltaic (PV) energy. PV energy is directly converting sunlight into electricity using the photoelectric effect. Since this has come into the limelight, the importance of having a reliable and accurate Total Cloud Cover (TCC) forecasting approach has attracted significant attention. Cloud cover plays a major role in variating solar irradiance. TCC is the fraction of the sky covered by all visible clouds. Variability in TCC contributes significantly to the final power output from PV energy sources. Therefore a reliable future TCC forecasting is required.

On another aspect, TCC is an important factor in predicting the weather. Clouds contribute to variability in the earth's surface temperature. If there is a high TCC, it can cause a drop in surface temperature. Apart from these, TCC directly or indirectly influences various fields such as agriculture, astronomy, tourism, and so on.

Therefore, this research focuses on the development of an efficient and reliable low-cost approach for short-term TCC forecasting. Hence our identified research questions are:

1) Can we estimate TCC using image processing techniques?
2) Can we generate a spatio-temporal time series dataset from cloud image sequences?
3) Can we use the generated time-series data for short-term TCC forecasting?
4) Can we develop a proof of concept for short-term TCC forecasting using satellite image sequences?

Motivated by its popularity and impact on the environment and society, we aim to develop an efficient and accurate future TCC forecasting system based on deep learning methods using satellite images. Specifically, a novel region based cloud segmentation approach is proposed for estimating TCC. A time-series dataset is also created by extracting image sequences from satellite images. The proposed system is composed of several key stages for TCC prediction. An edge detection function, the Otsu's method, is firstly used to calculate region based threshold values for cloud segmentation. The TCC information extracted from image sequences obtained from the segmented regional cloud cover is used to generate time-series data. The obtained time series data are subsequently used to train Long Short-Term Memory (LSTM) variant models for future TCC forecasting. We employ three LSTM variants, i.e. LSTM, bidirectional LSTM (BiLSTM) and Convolutional Neural Network-LSTM (CNN-LSTM), for performance comparison. The empirical results indicate that the proposed approach shows great efficiency in estimating future TCC from satellite images.

II. Related Work

Cloud information is usually associated with weather analysis and meteorological data. But as the main source of variability in solar irradiance [1], TCC is essential in solar irradiance estimation and PV energy generation prediction.

In the literature, TCC has been estimated with both ground sky camera images [2]–[5] and satellite images [6], [7]. Ground sky camera images are those that are captured by a special camera from ground level. Azhar et al. [4] have developed a cloud detection approach by calculating a threshold value for B-G color difference value in random pixels from a collected sky images data set. The effectiveness of B-G color difference approach has been proved in their study by a comparison carried out with R-G color difference approach which had been used in previous literature [8]. Besides the above, other edge detection methods such as the Otsu's thresholding method developed by Nobuyaski Otsu [9] has also been used

in previous studies for cloud identification in images [10], [11]. In addition, Zhen et al. [10] proposed a cloud detection approach using Total Sky Imager (TSI) images. TSI is a sky camera where the color images of the sky can be captured from the ground. Specifically, they first extracted the gray level matrices of the images. Then the Otsu's threshold method was used to identify the clouds from background. In addition to these traditional edge detection approaches, both algorithmic-based and machine learning-based approaches have been used for cloud identification in images. Kim et al. [2] and Lothon et al. [3] have developed algorithms to estimate TCC using sky camera images. Specifically, Kim et al. [2] have used a variable threshold value for red blue ratio, which has been determined from the frequency distribution of the green blue ratio. Their algorithm has resulted in low average RMSE values for cloud identification in ground sky images.

In recent studies, deep learning methods have been employed for TCC estimation. Xie et al. [5] have developed a deep CNN model, namely SegCloud, for the identification of clouds in ground sky camera images. SegCloud has resulted in an average segmentation accuracy rate of 96.24%, whereas the traditional R-B color difference threshold method has obtained a much lower accuracy rate of 81.17% for cloud segmentation in ground sky images. Similarly, deep CNN models have been used in a variety of other related studies for cloud detection in sky/satellite images [6], [7].

In most of the existing studies, extracted information from ground sky camera images and satellite images has been used for short-term forecasting of TCC. For long-term forecasting, development of numerical weather prediction (NWP) models has been exploited [12]. With the development of recurrent neural networks (RNNs) such as LSTM [13] and Gated Recurrent Unit (GRU) [14], the use of these deep neural network variants for TCC, PV energy, and solar irradiance forecasting has been extensively studied in recent years [15]–[19]. Specifically, the RNN and LSTM networks have feedback loops which allow them to maintain information in memory over the time. An LSTM unit as shown in Figure 1 in particular consists of three gates, i.e. input gate, output gate, forget gate, and a cell state memory. This has provided a solution to the RNN's vanishing gradient problem and enabled it to work well with time series data with long-term dependencies. As an example, Rajguruk et al. [20] identified that LSTM performs better than other RNNs for forecasting solar irradiance with time-series data. Another related study [19] from their team has also employed data extracted from ground sky camera images for solar irradiance forecasting using an LSTM network to indicate its efficiency. Their method was able to forecast solar irradiance with only 25.1% relative root mean square differences for 10-minute ahead forecasting. Prior work on video and image sequence analysis includes spatio-temporal feature extraction for generating numerical time-series datasets and dimension reduction in videos [21]–[36]. In particular, visual features like linear pixel distances [21]–[23] and color intensity values [24], [25], [37] have been adopted in existing video analysis studies.

By referring to the above existing literature, this research focuses on improving the existing state-of-the-art methods. For cloud segmentation, existing methods do not consider the illumination differences in an image due to sun's positions. This will inevitably affect edge detection threshold values for cloud segmentation. To address this challenging issue, we propose a novel region based threshold calculation and cloud segmentation approach. Specifically, a region based time series feature extraction strategy is employed for cloud segmentation and TCC estimation. The LSTM, BiLSTM and CNN-LSTM models are used for TCC forecasting. Subsequently, we compare the performance of the three deep networks, where the most performant models are used and tested for multi-step TCC forecasting.

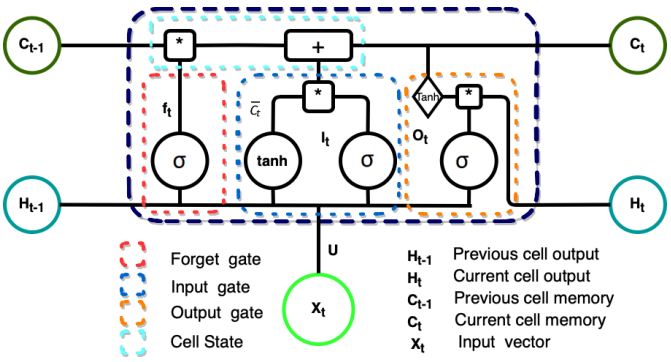

Fig. 1. LSTM unit (modified from [13])

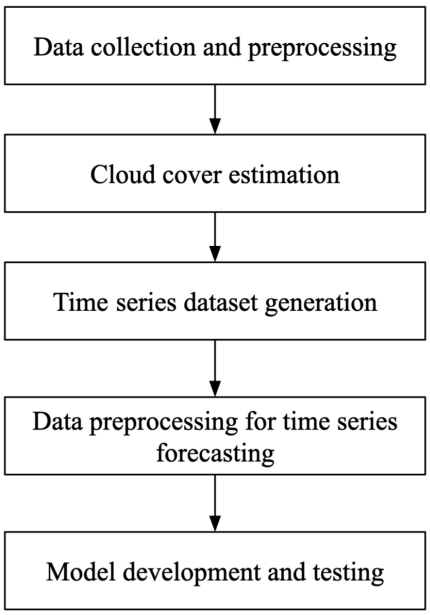

Fig. 2. The data flow of the proposed TCC forecasting system

III. The Proposed Methodology

In this research, first of all, an image sequence dataset was created from time-lapse videos captured from satellites. Precisely, time-lapse satellite videos were collected from publicly

available videos of Sat24.com website [38]. Each collected time lapse video is of 25 fps frame-rate. Each video consists of approximately 10-second video segment containing the day time video from sunrise to sunset. Each of these 10-second videos creates a sequence of images (approximately 250 images in total). For each image in a sequence of images, TCC was estimated with the proposed cloud segmentation approach based on the the Otsu's method. A time-series dataset has been subsequently generated using a new feature extraction method. The yielded time series dataset was then used for training and testing LSTM, BiLSTM aand CNN-LSTM respectively for future TCC forecasting. Figure 2 shows the detailed steps of the proposed system.

A. Total cloud cover estimation

In the collected satellite image sequence dataset, the earth's surface is in a darker color and clouds are in white or a comparatively lighter color. It was identified that, with a binary image conversion approach, pixels covered by clouds could be identified. First, each image was converted to a grayscale single-channel image. Next, a threshold value was applied to each image to convert the single-channel image to a binary one. For calculation of the threshold value per image, the Otsu's thresholding method was used. To be specific, the Otsu's method recursively searches for a threshold value that minimizes the intra-class variance score σ_ω^2, defined as the weighted sum of two variances as shown in (1).

$$\sigma_\omega^2 = \omega_b(t) * \sigma_b^2(t) + \omega_f(t) * \sigma_f^2(t) \tag{1}$$

where $\omega_b(t)$ and $\omega_f(t)$ are the probabilities of numbers of pixels for background and foreground classes at threshold t, respectively.

$$\omega_b(t) = \sum_{i=1}^{t} P(i) \tag{2}$$

$$\omega_f(t) = \sum_{i=t+1}^{l} P(i) \tag{3}$$

$\mu_b(t)$ and $\mu_f(t)$ are the mean values of background and foreground classes at the threshold t, respectively.

$$\mu_b(t) = \frac{\sum_{i=1}^{t} i * P(i)}{\omega_b(t)} \tag{4}$$

$$\mu_f(t) = \frac{\sum_{i=t+1}^{l} i * P(i)}{\omega_f(t)} \tag{5}$$

$\sigma_b^2(t)$ and $\sigma_f^2(t)$ are variances of background and foreground classes at threshold t, respectively.

$$\sigma_b^2(t) = \frac{\sum_{i=1}^{t} (1 - \mu_b(t))^2 * p(i)}{\omega_b(t)} \tag{6}$$

$$\sigma_f^2(t) = \frac{\sum_{i=t+1}^{l} (1 - \mu_f(t))^2 * p(i)}{\omega_f(t)} \tag{7}$$

From any generated binary image, TCC could be estimated by calculating cloud pixels in white. Figure 3 shows the comparison between an example image and the generated binary image with the Otsu's thresholding method.

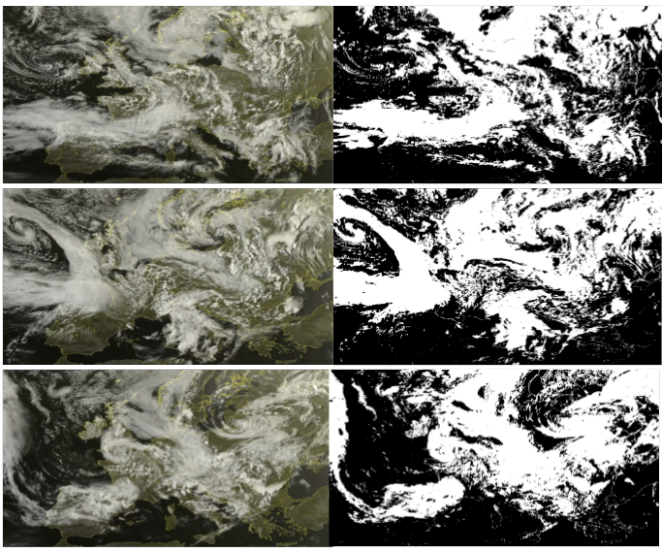

Fig. 3. Results with the Otsu's thresholding

Although cloud identification was satisfactory with the generated binary images as indicated in the above example, it was identified that illumination or the brightness of different regions of the image changes with time due to the movement of the sun. Therefore further investigation proved that the cloud cover identification could be improved by generating binary images using a region based Otsu's threshold value. Therefore, instead of applying the Otsu's method for the overall image directly, we divide each image into nine equal named regions as shown in Figure 4 where the Otsu's threshold values were calculated per named region. Figure 5 shows improvements with respect to the proposed region-based Otsu's thresholding than applying the thresholding method using the image threshold value at the image level. Figure 6 depicts the detailed steps for the proposed TCC prediction.

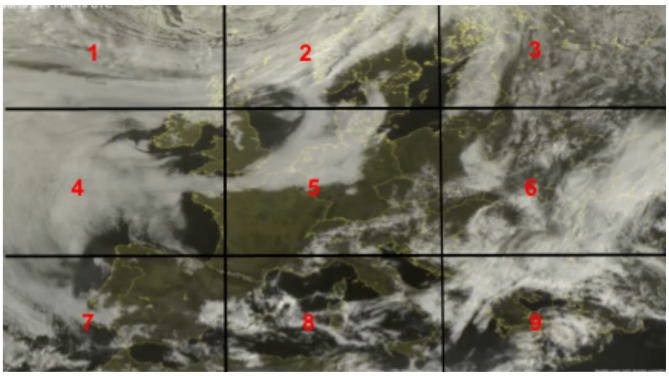

Fig. 4. The divided nine named regions

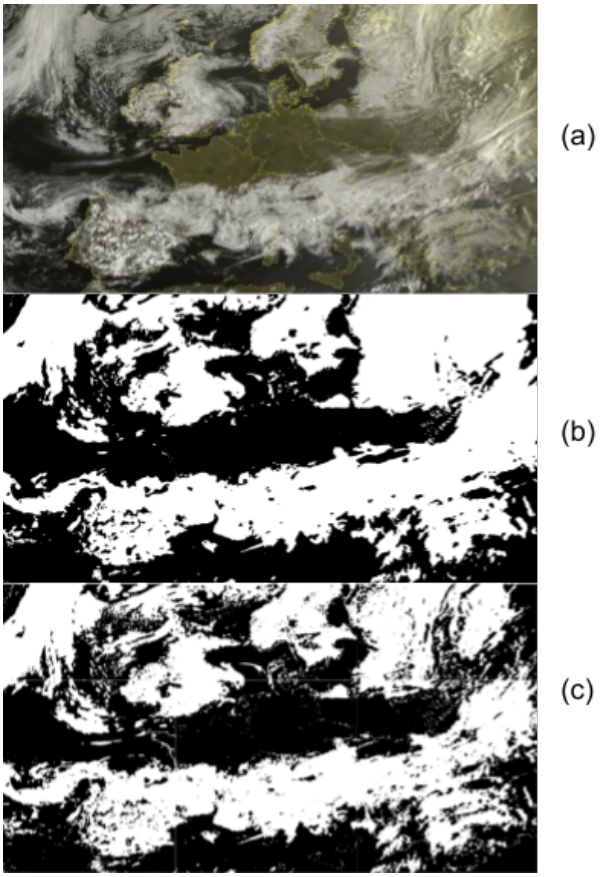

Fig. 5. Comparison between (a) the actual image, (b) per image Otsu's thresholding and (c) per named region Otsu's thresholding

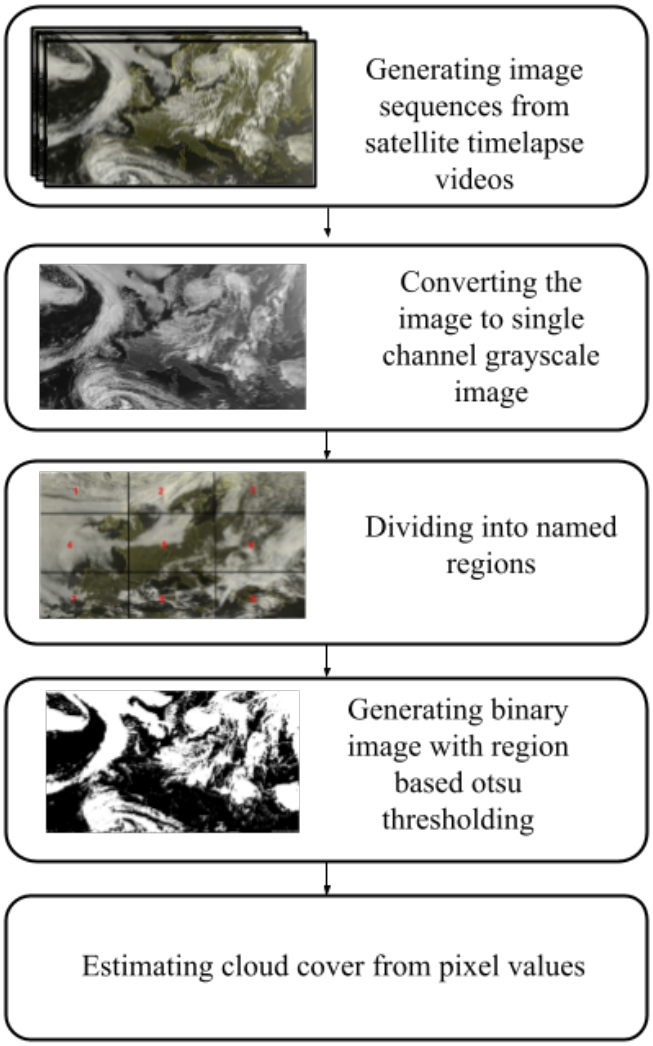

Fig. 6. The proposed approach for TCC estimation

B. Time-series dataset generation

By going through the collected image sequences, it was observed that short-term future TCC of a certain area is influenced by moving clouds in the surrounding areas. Therefore TCC values were estimated per named region (see Figure 4) assuming that the short-term future cloud coverage of a middle centred region will be mainly influenced by the current cloud coverage of the surrounding eight regions. It was also assumed that the visible sky area from a particular location on earth is denoted by a named region. For each named region, TCC was estimated by calculating the proportion of white pixels in the generated binary image.

By using these estimated TCC values, a time series dataset with nine features was generated for each image sequence. Each image sequence contains 250 images covering TCC variation throughout the daytime of a particular day. A dataset of 12,500 images is therefore generated for subsequent time series forecasting. The generated dataset consists of 112,500 estimated TCC values with a mean frequency score of 0.4420. To be specific, the dataset consists of frequency values between 0.01-0.86. The detailed frequency distribution is provided in Figure 7, while distribution measures of the dataset are also shown in Table I.

TABLE I
DISTRIBUTION MEASURES OF THE DATASET

Measure	*Value*
Range ($x_{max} - x_{min}$)	0.8500
InterQuartile Range (IQR)	0.2348
Mean	0.4420
Standard Deviation	0.1511
Coefficient of Variation	0.3420

C. Data pre-processing for time-series forecasting

The generated time-series data were then rearranged for time-series forecasting. The time-series data generated from 25 consecutive images were used as the input data and the time-series data generated from the next 10 consecutive images in the image sequence were considered as the target prediction. Data generated from 40 image sequences (i.e. 10,000 images) were considered as the training data and data generated from 10 image sequences (2,500 images) were employed as test data in this experiment.

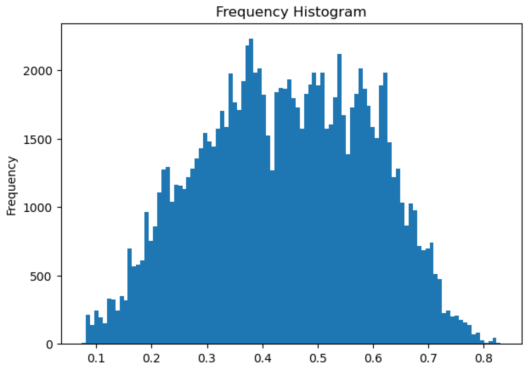

Fig. 7. Frequency distribution

D. Model development and testing

In this research, LSTM, BiLSTM and CNN-LSTM models are used to test and compare the results of time series forecasting. These deep networks are chosen owing to their impressive performances for time-series forecasting in existing studies [18]. We subsequently compare the performances of the three deep networks. The input shape of the networks is (25, 9) and the output data shape is (10, 9). Moreover, the model configurations of the aforementioned models are listed in Table II, Table III and Table IV respectively.

For all the trained models, the root mean squared error function is used as the loss function, while the "rmsprop" method is used as the optimizer. For evaluation and result comparison, a normalized root mean squared error (NRMSE) is calculated. To be specific, root mean squared error (RMSE) as defined in (8) is a standard metrics to measure the error of a model in quantitative data forecasting. It provides a measurement of the deviation between the predicted and the real values. The smaller the RMSE score, more accurate of the forecasting is. For this research, NRMSE as indicated in (9) was calculated by dividing the RMSE by the value range.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{N}} \tag{8}$$

where RMSE denotes the root mean square error, and N is the number of data points. In addition, x_i indicates the actual observations of time-series values, and $\hat{x}_i$ denotes the estimated time-series values.

$$NRMSE = \frac{RMSE}{x_{max} - x_{min}} \tag{9}$$

where NRMSE denotes the normalized root mean square error, while x_{max} and x_{min} indicate the maximum and minimum values of the dataset.

IV. Evaluation

For each neural network, the average RMSE scores were calculated against the validation dataset. To find the most suitable models for this task, a total of five rounds of testing were carried out with each proposed network. The BiLSTM model obtains the lowest average RMSE results. Figure 8, Figure 9 and Figure 10 show the training and testing RMSE score variations in accordance with different training epochs in LSTM, BiLSTM and CNN-LSTM networks, respectively. Table V shows the detailed training and testing RMSE scores for all the three models.

TABLE II
Model and Parameter Summary of LSTM Network

Layer	*Output Shape*	*Parameters*
LSTM	(None,25,200)	168000
Dropout	(None,25,200)	0
LSTM	(None,50)	50200
Dense	(None,90)	4590
Reshape	(None,10,9)	0
Total parameters		222790
Trainable parameters		222790

TABLE III
Model and Parameter Summary of Bi-directional LSTM Network

Layer	*Output Shape*	*Parameters*
Bi-directional LSTM	(None,25,100)	24000
Bi-directional LSTM	(None,100)	60400
Dense	(None,90)	9090
Reshape	(None,10,9)	0
Total parameters		93490
Trainable parameters		93490

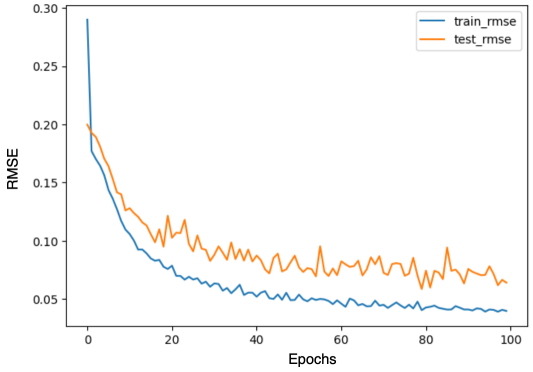

Fig. 8. RMSE variations with different training epoch settings in the LSTM network

As indicated in Table V, it is identified that the proposed BiLSTM model achieves the best performance with the lowest average RMSE and NRMSE scores. For instance, it obtains the lowest RMSE score of 0.0275 (i.e. a NRMSE score of 0.0323).

The best performant BiLSTM model with an RMSE score of 0.0275 is then tested with multi-step forecasting. First, the estimated TCC of succeeding 125 consecutive images is predicted using the estimated TCC of the first 25 consecutive images of an image sequence. In each image sequence generated from a

TABLE IV
MODEL AND PARAMETER SUMMARY OF CNN-LSTM NETWORK

Layer	*Output Shape*	*Parameters*
Time distributed conv 1D	(None,1,22,10)	370
Time distributed max pooling 1D	(None,1,11,10)	0
Time distributed flatten	(None,1,11,10)	0
LSTM	(None,100)	84400
LSTM	(None,1,20)	9680
Dense	(None,90)	1890
Reshape	(None,10,9)	0
Total parameters		96340
Trainable parameters		96340

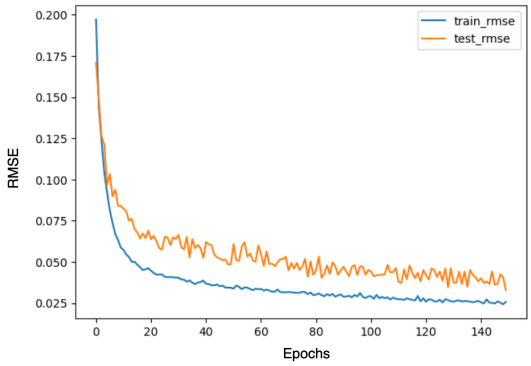

Fig. 9. RMSE variations with different training epoch settings in the BiLSTM network

time-lapse video, the full day time is covered with 250 images approximately. Therefore with 125 images forecasting from the first 25 images, the model's performance of forecasting TCC over the first half of daytime on a particular day can be evaluated.

The trained BiLSTM model is then used to forecast an estimated TCC of succeeding 125 images (from the first 25 images as input) with an average RMSE score of 0.0543 (i.e. a NRMSE result of 0.0639). The average RMSE result of 0.0823 (i.e. a NRMSE score of 0.0968) was achieved for TCC forecasting of succeeding 250 images with multi-step forecasting. The multi-step forecasting results for 10 days with respect to the BiLSTM model are shown in Figure 11.

V. DISCUSSION

In this research, satellite image sequences generated from recorded time-lapse videos were extracted for dataset construction. The yielded dataset was subsequently used to model the intended proof of concept for TCC prediction. Although the sufficient training is conducted using the dataset, lower RMSE values could have been achieved with a larger dataset. Also with a larger dataset, the numbers of named regions could have been increased.

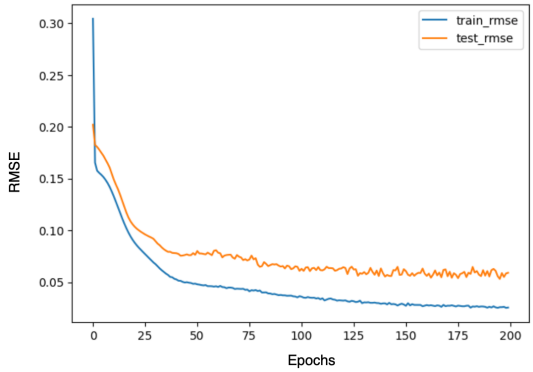

Fig. 10. RMSE variations with different training epoch settings in the CNN-LSTM network

TABLE V
FORECASTING ERROR RATES FOR DIFFERENT METHODS

Neural Network	*Training RMSE*	*Testing RMSE*
LSTM	0.0347	0.0563
	0.0345	0.0503
	0.0355	0.0474
	0.0352	0.0526
	0.0363	0.0487
Average RMSE	0.0352	0.0511
Average NRMSE	0.0414	0.0601
Bi-directional LSTM	0.0213	0.0282
	0.0217	0.0275
	0.0221	0.0292
	0.0223	0.0289
	0.0226	0.0279
Average RMSE	0.0220	0.0283
Average NRMSE	0.0259	0.0333
CNN LSTM	0.0293	0.0574
	0.0278	0.0550
	0.0284	0.0548
	0.0297	0.0532
	0.0284	0.0550
Average RMSE	0.0287	0.0551
Average NRMSE	0.0338	0.0648

In the proposed approach, the Otsu's thresholding method was used to identify the clouds in images for estimating the TCC. Threshold values have been generated per named region to capture illumination changes from one region to another. As indicated in the empirical results, the region based Otsu's thresholding works well in identifying clouds and sky/lands in the background. But this approach may result in inconsistencies in cloud identification when there is more noise or in the cases that the earth's surface is whiter (e.g. for the satellite images of polar regions). Therefore as a future implication, the use of a more complex approach for cloud identification in satellite images could be investigated.

The proposed approach uses image sequences and has divided satellite images into multiple regions assuming each region's future TCC is influenced by surrounding regions' current TCC. For each region, TCC is estimated in each image. With these extracted features, impressive RMSE scores

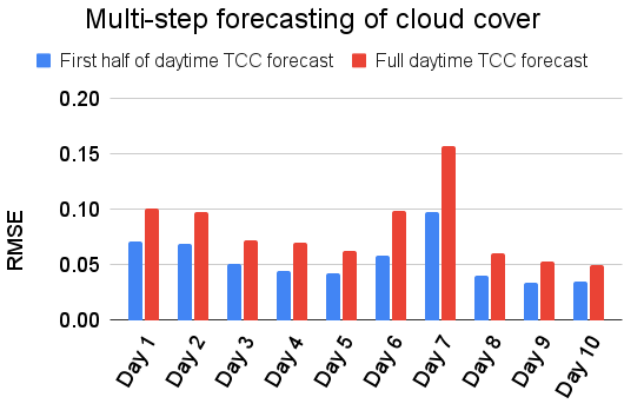

Fig. 11. Results of multi-step forecasting using the best performant BiLSTM network

have been achieved with all three proposed deep networks. In particular, the proposed BiLSTM model achieves the best average RMSE result of 0.0275 (i.e. the average NRMSE score of 0.0323) for forecasting the TCC of the succeeding 10 images from data extracted from 25 images. With multi-step forecasting, this best performant BiLSTM model has achieved an RMSE result of 0.0543 (a NRMSE score of 0.0639) in average for forecasting TCC over the first half of daytime on a particular day and 0.0823 RMSE in average (a NRMSE score of 0.0968) for forecasting TCC over a full daytime on a particular day. For this forecasting, the trained model only requires TCC data of 1/10-th of daytime on a particular day. Overall, the proposed region-based image segmentation based on the Otsu's method as well as the three proposed deep networks (LSTM, BiLSTM and CNN-LSTM) accounts for the superior performance of our proposed system for TCC forecasting.

VI. Conclusion

In this research, we have proposed region-based cloud segmentation using the Otsu's method and three deep networks for short-term TCC forecasting. The proposed system uses satellite image sequences to identify TCC variations and forecast future TCC from extracted regional features. Specifically, clouds in satellite images are effectively identified by applying the thresholding-based Otsu's method to the nine named regions. From visual observation of the segmented results, the region-based Otsu thresholding method has masked cloud cover better than the image-level Otsu's thresholding. Moreover, LSTM, BiLSTM and CNN-LSTM have produced competent RMSE results for TCC prediction based on time-series data extracted from the aforementioned regional features, which ascertains the efficiency of the proposed models for TCC forecasting.

Owing to the efficiency and light-weight characteristics of the proposed deep networks, our approach could be easily deployed to IoT devices to aid the general public in estimating future TCC in diverse real-world deployments. Furthermore, since our system does not require specialist machinery and equipment which may incur additional cost, the development of such a low-cost efficient method without additional hardware requirements could also open up future research in using the proposed method with mobile technologies, pertaining to energy efficiency and net-zero emission applications. In addition, hyper-parameter fine-tuning and deep architecture generation with attention mechanisms will also be exploited using evolutionary algorithms to further enhance performance [39]–[47].

References

[1] D. Matuszko, "Influence of the extent and genera of cloud cover on solar radiation intensity," *International Journal of Climatology*, vol. 32, no. 15, pp. 2403–2414, Dec. 2012. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/joc.2432

[2] B.-Y. Kim, J.-B. Jee, I.-S. Zo, and K.-T. Lee, "Cloud cover retrieved from skyviewer: A validation with human observations," *Asia-Pacific Journal of Atmospheric Sciences*, vol. 52, no. 1, pp. 1–10, Feb. 2016. [Online]. Available: http://link.springer.com/10.1007/s13143-015-0083-4

[3] M. Lothon, P. Barnéoud, O. Gabella, F. Lohou, S. Derrien, S. Rondi, M. Chiriaco, S. Bastin, J.-C. Dupont, M. Haeffelin, J. Badosa, N. Pascal, and N. Montoux, "ELIFAN, an algorithm for the estimation of cloud cover from sky imagers," *Atmospheric Measurement Techniques*, vol. 12, no. 10, pp. 5519–5534, Oct. 2019. [Online]. Available: https://amt.copernicus.org/articles/12/5519/2019/

[4] M. A. D. M. Azhar, N. S. A. Hamid, W. M. A. W. M. Kamil, and N. S. Mohamad, "Daytime Cloud Detection Method Using the All-Sky Imager over PERMATApintar Observatory," *Universe*, vol. 7, no. 2, p. 41, Feb. 2021. [Online]. Available: https://www.mdpi.com/2218-1997/7/2/41

[5] W. Xie, D. Liu, M. Yang, S. Chen, B. Wang, Z. Wang, Y. Xia, Y. Liu, Y. Wang, and C. Zhang, "SegCloud: a novel cloud image segmentation model using a deep convolutional neural network for ground-based all-sky-view camera observation," *Atmospheric Measurement Techniques*, vol. 13, no. 4, pp. 1953–1961, Apr. 2020. [Online]. Available: https://amt.copernicus.org/articles/13/1953/2020/

[6] M. Le Goff, J.-Y. Tourneret, H. Wendt, M. Ortner, and M. Spigai, "Deep Learning for Cloud Detection," in *8th International Conference of Pattern Recognition Systems (ICPRS 2017)*. Madrid, Spain: Institution of Engineering and Technology, 2017, pp. 10 (6 .)–10 (6 .). [Online]. Available: https://digital-library.theiet.org/content/conferences/10.1049/cp.2017.0139

[7] Z. Li, H. Shen, Q. Cheng, Y. Liu, S. You, and Z. He, "Deep learning based cloud detection for medium and high resolution remote sensing images of different sensors," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 150, pp. 197–212, Apr. 2019. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0924271619300565

[8] A. Heinle, A. Macke, and A. Srivastav, "Automatic cloud classification of whole sky images," *Atmospheric Measurement Techniques*, vol. 3, 05 2010.

[9] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, Jan. 1979. [Online]. Available: http://ieeexplore.ieee.org/document/4310076/

[10] Zhao Zhen, Bo Wang, Fei Wang, Jing Lu, Xiaoli Liu, Hui Ren, Chun Liu, Hongping Li, and Hongbin Sun, "Cloud identification model for sky images based on Otsu," in *International Conference on Renewable Power Generation (RPG 2015)*. Beijing, China: Institution of Engineering and Technology, 2015, pp. 5 .–5 . [Online]. Available: https://digital-library.theiet.org/content/conferences/10.1049/cp.2015.0521

[11] R. Chauvin, J. Nou, S. Thil, A. Traoré, and S. Grieu, "Cloud Detection Methodology Based on a Sky-imaging System," *Energy Procedia*, vol. 69, pp. 1970–1980, May 2015. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1876610215005044

[12] Baran, S. Lerch, M. El Ayari, and S. Baran, "Machine learning for total cloud cover prediction," *Neural Computing and Applications*, vol. 33, no. 7, pp. 2605–2620, Apr. 2021. [Online]. Available: https://link.springer.com/10.1007/s00521-020-05139-4

[13] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997. [Online]. Available: https://direct.mit.edu/neco/article/9/8/1735-1780/6109

[14] K. Chao, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation," 2014. [Online]. Available: https://arxiv.org/pdf/1406.1078v3.pdf

[15] A. Alzahrani, P. Shamsi, C. Dagli, and M. Ferdowsi, "Solar Irradiance Forecasting Using Deep Neural Networks," *Procedia Computer Science*, vol. 114, pp. 304–313, 2017. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1877050917318392

[16] F. Wang, Y. Yu, Z. Zhang, J. Li, Z. Zhen, and K. Li, "Wavelet Decomposition and Convolutional LSTM Networks Based Improved Deep Learning Model for Solar Irradiance Forecasting," *Applied Sciences*, vol. 8, no. 8, p. 1286, Aug. 2018. [Online]. Available: http://www.mdpi.com/2076-3417/8/8/1286

[17] M. Husein and I.-Y. Chung, "Day-Ahead Solar Irradiance Forecasting for Microgrids Using a Long Short-Term Memory Recurrent Neural Network: A Deep Learning Approach," *Energies*, vol. 12, no. 10, p. 1856, May 2019. [Online]. Available: https://www.mdpi.com/1996-1073/12/10/1856

[18] B. Brahma and R. Wadhvani, "Solar Irradiance Forecasting Based on Deep Learning Methodologies and Multi-Site Data," *Symmetry*, vol. 12, no. 11, p. 1830, Nov. 2020. [Online]. Available: https://www.mdpi.com/2073-8994/12/11/1830

[19] R. A. Rajagukguk, R. Kamil, and H.-J. Lee, "A Deep Learning Model to Forecast Solar Irradiance Using a Sky Camera," *Applied Sciences*, vol. 11, no. 11, p. 5049, May 2021. [Online]. Available: https://www.mdpi.com/2076-3417/11/11/5049

[20] R. A. Rajagukguk, R. A. A. Ramadhan, and H.-J. Lee, "A Review on Deep Learning Models for Forecasting Time Series Data of Solar Irradiance and Photovoltaic Power," *Energies*, vol. 13, no. 24, p. 6623, Dec. 2020. [Online]. Available: https://www.mdpi.com/1996-1073/13/24/6623

[21] I. Bandara and B. Bačić, "Strokes classification in cricket batting videos," in *2020 5th International Conference on Innovative Technologies in Intelligent Systems and Industrial Applications (CITISIA)*, 2020, pp. 1–6.

[22] B. Bačić and I. Bandara, "Tennis strokes recognition from generated stick figure video overlays," in *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 4: VISAPP,*, INSTICC. SciTePress, 2022, pp. 397–404.

[23] G. Khan, A. Siddiqi, M. U. Khan, S. Wahla, and S. Samyan, "Geometric positions and optical flow based emotion detection using mlp and reduced dimensions," *IET Image Processing*, vol. 13, 03 2019.

[24] J. K and S. Mary Idicula, "Enhanced video classification system using a block-based motion vector," *Information*, vol. 11, no. 11, 2020. [Online]. Available: https://www.mdpi.com/2078-2489/11/11/499

[25] A. Sharma and A. K. Singh, "Color difference histogram for feature extraction in video retrieval," 2015.

[26] T. Y. Tan, L. Zhang, and C. P. Lim, "Intelligent skin cancer diagnosis using improved particle swarm optimization and deep learning models," *Applied Soft Computing*, vol. 84, p. 105725, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S156849461930506X

[27] B. Fielding, T. Lawrence, and L. Zhang, "Evolving and ensembling deep cnn architectures for image classification," Jul. 2019, 2019 International Joint Conference on Neural Networks, IJCNN 2019 ; Conference date: 14-07-2019 Through 19-07-2019. [Online]. Available: https://www.ijcnn.org/

[28] T. Y. Tan, L. Zhang, C. P. Lim, B. Fielding, Y. Yu, and E. Anderson, "Evolving ensemble models for image segmentation using enhanced particle swarm optimization," *IEEE Access*, vol. 7, pp. 34 004–34 019, 2019.

[29] P. Kinghorn, L. Zhang, and L. Shao, "A hierarchical and regional deep learning architecture for image description generation," *Pattern Recognition Letters*, vol. 119, pp. 77–85, 2019, deep Learning for Pattern Recognition. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167865517303240

[30] H. Xie, L. Zhang, and C. P. Lim, "Evolving cnn-lstm models for time series prediction using enhanced grey wolf optimizer," *IEEE Access*, vol. 8, pp. 161 519–161 541, 2020.

[31] T. Y. Tan, L. Zhang, and C. P. Lim, "Adaptive melanoma diagnosis using evolving clustering, ensemble and deep neural networks," *Knowledge-Based Systems*, vol. 187, p. 104807, 2020.

[32] L. Zhang and C. P. Lim, "Intelligent optic disc segmentation using improved particle swarm optimization and evolving ensemble models," *Applied Soft Computing*, vol. 92, p. 106328, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494620302684

[33] L. Zhang, C. P. Lim, and Y. Yu, "Intelligent human action recognition using an ensemble model of evolving deep networks with swarm-based optimization," *Knowledge-Based Systems*, vol. 220, p. 106918, 2021.

[34] Y. Zhang, L. Zhang, and M. A. Hossain, "Adaptive 3d facial action intensity estimation and emotion recognition," *Expert systems with applications*, vol. 42, no. 3, pp. 1446–1464, 2015.

[35] Y. Zhang, L. Zhang, S. C. Neoh, K. Mistry, and M. A. Hossain, "Intelligent affect regression for bodily expressions using hybrid particle swarm optimization and adaptive ensembles," *Expert Systems with Applications*, vol. 42, no. 22, pp. 8678–8697, 2015.

[36] T. Lawrence and L. Zhang, "Iotnet: An efficient and accurate convolutional neural network for iot devices," *Sensors*, vol. 19, no. 24, p. 5541, 2019.

[37] S. Cvetkovic, M. Jelenkovic, and S. Nikolic, "Video summarization using color features and efficient adaptive threshold technique," *Przeglad Elektrotechniczny*, 2013.

[38] "Weather Europe, Satellite Weather Europe, Weather Forecast, Rainfall, Clouds, Sun in Europe - Source: SAT24.com," 2021. [Online]. Available: https://www.sat24.com

[39] S. Neoh, L. Zhang, K. Mistry, M. Hossain, C. Lim, N. Aslam, and P. Kinghorn, "Intelligent facial emotion recognition using a layered encoding cascade optimization model," *Applied Soft Computing*, vol. 34, pp. 72–93, Sep. 2015.

[40] K. Mistry, L. Zhang, S. C. Neoh, C. P. Lim, and B. Fielding, "A micro-ga embedded pso feature selection approach to intelligent facial emotion recognition," *IEEE Transactions on Cybernetics*, vol. 47, no. 6, pp. 1496–1509, 2017.

[41] P. Kinghorn, L. Zhang, and L. Shao, "A region-based image caption generator with refined descriptions," *Neurocomputing*, vol. 272, pp. 416–424, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231217312511

[42] L. Zhang, K. Mistry, C. P. Lim, and S. C. Neoh, "Feature selection using firefly optimization for classification and regression models," *Decision Support Systems*, vol. 106, pp. 64–85, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167923617302208

[43] C. Wall, L. Zhang, Y. Yu, and K. Mistry, "Deep recurrent neural networks with attention mechanisms for respiratory anomaly classification," in *2021 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2021, pp. 1–8.

[44] T. Lawrence, L. Zhang, C. P. Lim, and E.-J. Phillips, "Particle swarm optimization for automatically evolving convolutional neural networks for image classification," *IEEE Access*, vol. 9, pp. 14 369–14 386, 2021.

[45] T. Lawrence, L. Zhang, K. Rogage, and C. P. Lim, "Evolving deep architecture generation with residual connections for image classification using particle swarm optimization," *Sensors*, vol. 21, no. 23, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/23/7936

[46] L. Zhang, C. P. Lim, Y. Yu, and M. Jiang, "Sound classification using evolving ensemble models and particle swarm optimization," *Applied Soft Computing*, vol. 116, p. 108322, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494621011157

[47] D. Pandit, L. Zhang, S. Chattopadhyay, C. P. Lim, and C. Liu, "A scattering and repulsive swarm intelligence algorithm for solving global optimization problems," *Knowledge-Based Systems*, vol. 156, pp. 12–42, 2018.